DOCUMENT RESUME

ED 214 882 SP 019 819

AUTHOR Sirotnik, Kenneth A.
TITLE Assessing Attitudinal Congruency: A Case for Absolute (As Well As Relative) Indices. A Study of Schooling in the United States. Technical Report Series, No. 12.
INSTITUTION California Univ., Los Angeles. Graduate School of Education
SPONS AGENCY Institute for Development of Educational Activities, Dayton, Ohio.
PUB DATE 80
NOTE 18p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Affective Measures; *Attitude Measures; *Behavioral Science Research; Comparative Analysis; *Congruence (Psychology); *Correlation; Decision Making; Predictive Measurement; *Projective Measures; Research Methodology; Research Problems

ABSTRACT

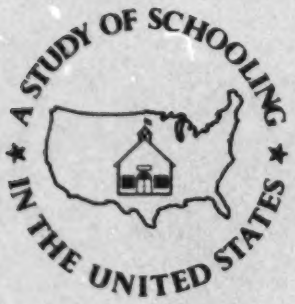
Two analytic models, each with different implications for data analysis, have been used in behavioral science research for assessing attitudinal congruency. The term relative congruence applies to the family of agreement indices which, although functions of the absolute index, are formulated relative to chance expectations. The term absolute congruency applies to the simple observed proportion of agreement between respondent pairs. Although behavioral research literature points to the almost universal use of relative congruency, its use is sensible only in those circumstances where the issue of consistency is primarily one of predictability or joint attitudinal disposition. However, if the researcher's interest focuses on how much congruency exists, the absolute index (observed congruence percentage) is a more relevant statistic, especially in comparative studies. An argument can also be made for the coexistence of both measures, each focused on statistical and substantive issues which are not necessarily the same. In some problem solving instances, planners not only must rely on predicting a certain outcome based on limited attitudinal data, but also must consider the extent to which subjects hold the same attitudes. The intended use of the results should be the determinant of the analytic model(s) used. (FG)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED214882

A STUDY OF SCHOOLING

Technical Report Series



Distributed by the Laboratory in
School and Community Education

Graduate School of Education
University of California
Los Angeles, California 90024

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

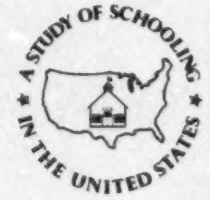
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

John I. Goodlad

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

SP 019 819



SUPPORTED BY

The Danforth Foundation
International Paper Company Foundation
The JDR 3rd Fund
Martha Holden Jennings Foundation
Charles F. Kettering Foundation
Lilly Endowment, Inc.
Charles Stewart Mott Foundation
National Institute of Education
The Needmor Fund
Pedamorphosis, Inc.
The Rockefeller Foundation
The Spencer Foundation
United States Office of Education

Conducted under the auspices of the Institute for
Development of Educational Activities (/I/D/E/A/)
A Division of the Charles F. Kettering Foundation

John I. Goodlad, Principal Investigator

ASSESSING ATTITUDINAL CONGRUENCY: A CASE FOR ABSOLUTE (AS WELL AS RELATIVE) INDICES

Kenneth A. Sirotnik

Technical Report No. 12

1980

A Study of Schooling is based upon the assumption that improving schools requires knowing what is happening in and around them. A comprehensive data-base of contextual information was obtained from students, teachers, administrators, parents and observers at all grade levels in thirty-eight elementary and secondary purposively sampled schools. *It is strongly recommended that readers of any technical report in this series first read Technical Report No. 1 which outlines the details, scope and limitations of the Study as a whole.*

It must be understood that this series of technical reports does not constitute the Study. Some reports are highly specific "molecular" inquiries while others take a more "molar" view across data sources, schooling levels, etc. Some reports are more methodological in nature arising out of issues in data analysis. Many of the reports quite naturally overlap in data analysed and interpretations rendered. Some authors have approached their task as consisting mostly of data description with little discussion beyond the presentation of the data. Others have ventured further into the realm of interpretation and speculation. *It must be further understood that data-based inferences can and do differ among researchers who come at the data from differing points-of-view.* Authors, therefore, are duly acknowledged for each report and are responsible for the material presented therein.

Assessing Attitudinal Congruency: A Case for Absolute (as well as Relative) Indices

Introduction

Looking through most of the more popular textbooks on research methods (including those on statistics) used in the behavioral sciences, one will be unable to find any presentation focused on the measurement of attitude congruence. Yet, the notion of agreement (or disagreement) between matched-pairs of some type or another has been widely researched in many areas, e.g., parents and children; husbands and wives; teachers and students; etc. Perhaps it has just been assumed that assessing this kind of agreement or congruency is merely an application of well-known methods such as correlational analysis and comparison of means.[1] But it will be argued here that questions of attitude congruency can be of at least two distinct types (absolute and/or relative) with attendant differences in analytic models and, therefore, different implications for data analysis.

The term absolute congruency applies to the simple observed proportion of agreement between respondent pairs. The term relative congruency applies to the family of agreement indices which, although functions of the absolute index, are formulated relative to chance expectations. It should be noted that differentiating between these indices is a general problem pertaining to any form of cross-classification for the purpose of measuring agreement. For example, the issue raised herein has recently been confronted and similarly resolved by Kane and Brennan (1980) in the context of criterion-referenced testing and indexing the reliability of mastery-nonmastery decisions.

In this paper, the problem will be addressed in the context of assessing attitudinal congruency. Although the absolute and relative congruency indices originally emerged in the literature on nominal scale agreement, they are

equally suitable for attitude measures on ordinal (and often assumed equal interval) scales. By way of example, it will be shown how very different quantitative impressions of the same data set can be created by each of these congruency approaches. Moreover, either one (or both) of the resultant congruency index types will be relevant depending upon the purpose(s) behind the research question(s). The focus, however, will be on the absolute index since it has been almost totally subordinated by those indexes derived from the relative or "corrected-for-chance" perspective.

Review of the Basic Design

The most complete description of response congruency obtains when the responses of the members of the pairs in question are crosstabulated. For example, Table 1 presents such an analysis for a typical Likert-type attitude survey item in both the parent and secondary student questionnaires

- - - - - - - - - - - - - -

Insert Table 1 about here

- - - - - - - - - - - - - -

used in A Study of Schooling.[2] The item pertains to one's willingness to be bused (student form) or to have one's child bused (parent form) to achieve school desegregation. Both parents and their children responded to this issue ("personal commitment to busing") on a 4-point agreement scale (4/3 = strongly/mildly agree; 2/1 = mildly/strongly disagree). Adding up the overall percentages (third number in each cell) in the main diagonal cells, we see that nearly half (47%) of the parent-child pairs respond identically to the issue. This is an index of absolute congruency.

The question of congruency, however, can also be approached from a relative standpoint by comparing what is observed (i.e. the absolute measure) with what is expected from the marginal response distributions of parents and

students separately. The fourth number in each cell of Table 1 is the expected percentage following ordinary chi-square procedures for analysing contingency tables. The absolute congruency percentage is, therefore, only 7% above the expected congruency of 40%.

Congruency indices can be briefly summarized by using the following notational system for any crosstabulation:

p_{ij} = cell proportion for row i and column j

$p_{i.}$, $p_{.j}$ = marginal proportions for row i ($\sum_j p_{ij}$) and column j ($\sum_i p_{ij}$) respectively

$p_{..} = \sum_i \sum_j p_{ij} = 1$ by definition

The observed proportion or absolute index of congruency is then $p_o = \sum_{i=j} p_{ij}$. Denoting the expected congruency proportion as p_e and the ideal congruency proportion as p_I, relative congruency indices proposed in the literature have the following form:

$$\text{Index} = \frac{p_o - p_e}{p_I - p_e} \qquad (1)$$

That is, they are ratios of the amount of observed congruency beyond that expected to that amount possible under idealized circumstances.

Variants of this index depend upon how the p_I and p_e are defined. The most common definition of p_I is 1, i.e., perfect congruency. In this case, Krippendorff (1970) has shown that (1) defines Scott's (1955) congruency index π when $p_e = \sum_{i=j} \frac{(p_{i.} + p_{.j})^2}{2}$ and Cohen's (1960) congruency index κ when $p_e = \sum_{i=j} p_{i.}\, p_{.j}$.[3] Cohen (1968) has also developed a weighted version of kappa by including multiplicative constants w_{ij} in a formula equivalent to (1) but written in terms of the incongruency cells p_{ij} $(i \neq j)$:

$$\kappa_w = 1 - \frac{\sum_{i \neq j} \sum w_{ij}\, p_{ij}}{\sum_{i \neq j} \sum w_{ij}\, p_{i.}\, p_{.j}}. \qquad (2)$$

It is in this formulation that the Pearson correlation coefficient r gains redeeming value as a congruency index. Cohen (1968) has shown that if (a) the p_{ij}'s are weighted by the square of their distance from the congruency diagonal (i.e., when $w_{ij} = |i - j|^2$) and (b) the marginal distributions are identical (i.e., $p_{i.} = p_{.j}$), then $\kappa \equiv r$. When $p_{i.} \neq p_{.j}$, $\kappa < r$, but usually the difference is not very great. Thus, unless a highly unlikely response pattern like the following occurred,

Parent Response	Student Response 1	2	3	4
1	0	X	0	0
2	0	0	X	0
3	0	0	0	X
4	0	0	0	0

where $\kappa = 0$ and $r = 1$, the correlation coefficient becomes a viable index of relative congruency for items with limited scale ranges.

A less common definition of p_I is the maximum observed congruency proportion possible given the marginal distributions. It is well known that indices based upon contingency tables are limited in value when the $p_{i.} \neq p_{.j}$. The maximum observed congruence proportion (p_{max}), for example, is limited to $\sum_{i=j} \text{Min}\,(p_{i.}, p_{.j})$ and is 1 only when $p_{i.} = p_{.j}$. Probably the first researcher to propose this variant of the index given by (1) above, Benini (1901), reasoned that p_{max} (not 1) was the ideal congruency. Benini's formula, then, was equivalent to (1) with $p_I = p_{max}$ and p_e computed as in the kappa coefficient. A little bit of algebraic manipulation shows that this variant is simply κ/κ_{max} or, in the dichotomous case, the well-known coefficient ϕ/ϕ_{max} (or nearly so in the case of $p_{i.} \neq p_{.j}$).

Absolute vs Relative Indices

Clearly, very different impressions of attitude congruency are conveyed by absolute and relative indices. With little loss in generality, the issue is nicely illustrated in 2 x 2 crosstabulations which highlight the difference between joint and marginal responses. Compare, for example, the hypothetical outcomes A and B in Table 2. Both have perfect relative indices (κ) of 1 and perfect

- - - - - - - - - - - -

Insert Table 2 about here

- - - - - - - - - - - - -

absolute indices of 100%, yet outcome B shows only an 18% difference between observed and expected percentages in contrast to the maximum possible 50% evidenced in outcome A. Outcome C has nearly the same absolute index and $p_o - p_e$ difference as outcome B; yet kappa is a mere .46 because of the slight shift of 6% to the incongruous cells. However, the main effect of 90% congruity in the (+,+) cell is still identical in both outcomes B and C. Finally, outcome D has a higher κ than outcome C; yet, 18% fewer cases are congruent in outcome D. Moreover, outcome D evidences the maximum possible congruence given the unequal marginal distributions, i.e., $\kappa/\kappa_{max} = 1$ for this outcome.

So which outcomes evidence more and less attitude congruence? It would appear that either index (absolute or relative) can be misleading. The argument here is for the coexistence of both measures, each focused on statistical (and substantive) issues which are not necessarily the same. The absolute index tells us how much or what amount of congruency, in fact, exists. It would be hard to argue that more matched-pairs in outcome D selected the same response positions than those in outcome C. The relative index reflects the liklihood of the observed congruency being a result of joint attitudinal dispositions on the part of the members of each pair. It would not be hard to argue that although congruency is lower in

outcome D, there is greater evidence to suggest that it is more a function of the paired unit than is the higher congruency in outcome C. Take the parent-child pairs as an example: a higher κ is more supportive of a hypothesis regarding the familial unit, e.g., one of shared values, generational transmission of values, etc. and, hence, of the predictability of parent and child responses, one from the other.

Given this difference in interpretation, it is easy to see why relative congruency indices have received more attention. The congruency problem has almost always been viewed in terms of interrater (or inter-observer) reliability for 2 fixed raters (observers). For example, instead of parent-child matched-pairs ("families") being the unit of replication as in the present formulation, objects (stimuli, etc.) constitute the replications in the 2-rater reliability problem. Light (1971) has noted that so long as the replications can be treated independently, both the interrater and matched-pair paradigms are statistically analogous. Importantly, however, the same type of research question (and, thus, the same statistical index) is not necessarily relevant in both problems.

In the interrater paradigm, the agreement percentage corrected for chance expectation has usually been viewed as more relevant than the agreement percentage itself, insofar as predictability from one rater to another is central to the reliability issue. This parallels Kane and Brennan's (1980) characterization of the relative index as estimating the "contribution of the test to the dependability of the decision procedure" in domain-referenced testing. Moreover, these authors also note the utility of the absolute index as an estimate of the dependability of the decision procedure itself.

In the matched-pair paradigm, the consequences of differentiating between absolute and relative indices is nicely illustrated when one is attempting a comparative study of attitude congruence. To keep matters simple, the data in Table 1 have been dichotomized and presented separately in Table 3 for the three major race/ethnicity groups in the Study

- - - - - - - - - - - - - -

Insert Table 3 about here

- - - - - - - - - - - - - -

of Schooling sample. None of the kappa coefficients are particularly impressive, although the .17 coefficient for the black parent-child pairs is at least non-zero. Does this mean that the black pairs are more attitudinally congruent than the white or brown pairs? Yes, in the sense that we have more evidence to believe that the observed congruency of 58% is due to something shared or transmitted in the familial unit. But no, in the sense that 73% of the white parents and their children demonstrate the same attitudinal response tendencies. To be sure, these same tendencies would have been evidenced had the parents and students been randomly selected and paired (so long as the same marginals were obtained). The data simply do not permit us to assume anything operating other than similar marginal attitudes.

The issue, then, boils down to the intended use of the results. For example, why would one do a parent-child study on attitude congruence? One possible rationale (see Sirotnik, 1981) is based on a decision-making/problem-solving model used in the community school approach. Such an approach requires a "meeting-of-the-minds" or, at least, a common understanding of the attitudinal differences among the critical participants. From this perspective, being able to predict the child's attitude from the parent's attitude--and vice-versa--is not so important as determining to what extent they hold the same attitudes. In effect, the observed percentages for the three groups in

Table 3 provide the most insightful data--not just the congruency percentages, but the pattern of percentages in the incongruous cells as well as the differences in marginal percentages. Again, the utility of relative indices like kappa is not being discounted; but, rather, emphasis is being placed upon the coordination of their use with specified purposes. (Note that, when appropriate, kappa can be partialized, pinpointing the congruency cell with more or less observed congruency beyond that expected.[4] Note, also, that phi coefficients were calculated in Table 3, mainly for the purpose of demonstrating their near equivalence to κ in spite of unequal marginal distributions.)

A more complete description of the comparative results in Table 3, then, is as follows: White parent-child pairs show the most attitude congruency (73%) on the busing issue, primarily due to their overwhelming response tendencies for <u>disagreement</u> (78% for children and 87% for parents). As a result, little association exists between parent and child responses ($\kappa=\phi=.08$), although some degree of predicability emerges in childrens' responses for agreeing parents ($\kappa_{row_1} = .12$). Mexican-American pairs evidence the least congruency (49%) and the most overall discrepency between parent and child attitudes (31% and 54% agreement for parents and children respectively). Virtually no association exists between parent and child attitudes for this sub-sample. Finally, black parent-child pairs evidence the most predictability in attitudinal response ($\kappa=\phi = .17$) with a moderate amount of congruency (58%) and nearly the same overall parent and child attitudinal response tendencies (approximately 50%).

<u>Summary</u>

What literature exists on the methodology for assessing attitudinal congruency points almost universally to the use of relative (correlation or correlation-like) coefficients. It has been attempted here to show with examples

how this procedure is sensible only in those circumstances where the issue of congruency is primarily one of predictability or joint attitudinal disposition. If, however, the researcher's interest focuses primarily on "how much" congruency exists--a perfectly legitimate question in many applied settings--the absolute index (observed congruency percentage) is a more relevant statistic, especially in comparative studies.

FOOTNOTES

1. Much of the literature on parent-child attitude studies has been reviewed elsewhere (Sirotnik, 1981) wherein three distinct analytic methods seemed to be in use equally often for assessing attitudinal congruency: (a) Pearson product-moment correlation (b) paired-comparison of means and (c) comparison of independent group means. Not only can these methods lead to non-convergent interpretations but method (b) is not necessarily sensitive to congruency since the algebraic sign of differences confounds the comparison; and method (c) is clearly inappropriate given the dependent nature of paired data.

2. More information on the background and current activities of A Study of Schooling can be found in the series of four articles published in the Phi Delta Kappan beginning with Goodlad, Sirotnik and Overman (1979).

3. Coefficient π assumes a single underlying response distribution estimated by averaging the marginal distributions for each respondent. Coefficient κ assumes two underlying distributions, each estimated by the two marginal distributions, i.e., the usual chi square approach, which, as argued by Cohen (1960), represents the reality of the response tendencies better than any assumption regarding these tendencies.

4. Partial kappa for any row i is computed using the following formula:

$$\kappa_{row_i} = \frac{p_{ii}/p_{i.} - p_{.i}}{1 - p_{.i}}$$

(See Coleman, 1966 and Light, 1971.)

REFERENCES

1. Benini, R. Principii di demografia. Mannual: Barbera di Science Giuridiche Social: e Politiche, 1901, No. 29.

2. Cohen, J.A. Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. Psychological Bulletin, 1968, 70, 213-220.

3. Cohen, J.A. A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 1960, 20, 37-46

4. Coleman, J.S. Measures of concordance or consensus between members of social groups. Unpublished manuscript, Johns Hopkins University. 1966.

5. Goodlad, J.I., Sirotnik, K.A., & Overman, B.C. A study of schooling: An overview. Phi Delta Kappan, 1979, 61, 176-178.

6. Kane, M.T. & Brennan, R.L. Agreement coefficients as indices of dependability for domain-referenced tests. Applied Psychological Measurement, 1980, 4, 105-126.

7. Krippendorff, K. Bivariate agreement coefficients for reliability data. In Borgatta, E.F. (Ed.) Sociological Methodology. San Francisco: Josey-Bass, Inc., 1970.

8. Light, R.J. Measures of response agreement for qualitative data: Some generalizations and alternatives. Psychological Bulletin, 1971, 76, 365-377.

9. Scott, W.A. Reliability of content analysis: The case of nominal scaling. Public Opinion Quarterly, 1955, 19, 321-325.

10. Sirotnik, K.A. Parents and their children: A study of congruence on attitudes about school. Study of Schooling, Technical Report No. 13, 1981.

TABLE 1

Crosstabulation of Parent and Child Attitudinal Responses
On a 4-Point Likert-Type Item

Parent Response	Child Response 1	2	3	4	Totals
1	1631[a]	517	442	254	2844
	57.3[b]	18.2	15.5	8.9	100.0
	38.9[c]	12.3	10.6	6.1	67.9
	(34.4)[d]	(13.6)	(11.9)	(8.0)	
2	221	121	103	54	499
	44.3	24.2	20.6	10.8	100.0
	5.3	2.9	2.5	1.3	11.9
	(6.0)	(2.4)	(2.1)	(1.4)	
3	173	125	105	79	482
	35.9	25.9	21.8	16.4	100.0
	4.1	3.0	2.5	1.9	11.5
	(5.8)	(2.3)	(2.0)	(1.4)	
4	95	79	83	106	363
	26.2	21.8	22.9	29.2	100.0
	2.3	1.9	2.0	2.5	8.7
	(4.4)	(1.7)	(1.5)	(1.0)	
Totals	2120	842	733	493	4188
	50.6	20.1	17.5	11.8	100.0

p_o = .468 κ = .12

p_e = .398 r = .23

a cell frequencies

b percentages of row totals

c percentages of table total

d expected cell percentages (of table total)

TABLE 2

Hypothetical Outcomes for
Matched-Pair 2 x 2 Crosstabulations

(cells contain observed & expected proportions p_{ij} as defined in narrative)

	Outcome	+	−		P_o	P_e	(P_o-P_e)	κ
A	+	.50 (.25)	0 (.25)	.50	1.00	.50	.50	1
	−	0 (.25)	.50 (.25)	.50				
		.50	.50	1				
B	+	.90 (.81)	0 (.09)	.90	1.00	.82	.18	1
	−	0 (.09)	.10 (.01)	.10				
		.90	.10	1				
C	+	.90 (.87)	.04 (.06)	.94	.93	.87	.06	.46
	−	.03 (.06)	.03 0	.06				
		.93	.07	1				
D	+	.35 (.21)	.25 (.39)	.60	.75	.47	.28	.53
	−	0 (.14)	.40 (.26)	.40				
		.35	.65	1				

TABLE 3

Crosstabulation of Dichotomized
Parent and Child Responses in Table 1
For Each of Three Race/Ethnicity Groups

Race/ Ethnicity	Parent Response	Child Response: Agree	Child Response: Disagree	Totals	Statistics[d]		
White	Agree	119[a]	262	381	p_o = 73.0	p_e = 70.6	
		31.2[b]	68.8	100.0			
		4.0[c]	8.7	12.7	κ = .08	ϕ = .08	κ/κ_{max} = .12
	Disagree	550	2071	2621	κ_{row_1} = .12	κ_{row_2} = .06	
		21.0	79.0	100.0			
		18.3	69.0	87.3			
	Totals	669	2333	3002			
		22.3	77.7	100.0			
Black	Agree	171	135	306	p_o = 58.3	p_e = 49.9	
		55.9	44.1	100.0			
		29.4	23.2	52.7	κ = .17	ϕ = .17	κ/κ_{max} = .19
	Disagree	107	168	275	κ_{row_1} = .15	κ_{row_2} = .19	
		38.9	61.1	100.0			
		18.4	28.9	47.3			
	Totals	278	303	581			
		47.8	52.2	100.0			
Mexican-American	Agree	64	52	116	p_o = 48.9	p_e = 48.3	
		55.2	44.8	100.0			
		17.1	13.9	31.0	κ = .01	ϕ = .01	κ/κ_{max} = .02
	Disagree	139	119	258	κ_{row_1} = .02	κ_{row_2} = .01	
		53.9	46.1	100.0			
		37.2	31.8	69.0			
	Totals	203	171	374			
		54.3	45.7	100.0			

a Cell frequency
b Percentage of row total
c Percentage of table total
d p_o, p_e, κ, ϕ, κ_{max} as defined in narrative;
κ_{row_1} and κ_{row_2} are partial kappa coefficients by row.